Legal Theory (2025), 31: 4, 293–315
doi:10.1017/S1352325226100810

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Disparate Treatment and Discriminatory Harm

Deborah Hellman[1] and Lily Hu[2]

[1]School of Law, University of Virginia, Charlottesville, VA, USA and [2]Department of Philosophy, Yale University, New Haven, CT, USA
Corresponding author: Deborah Hellman; Email: dhellman@law.virginia.edu

(Received 21 November 2024; Revised 05 August 2025; Accepted 30 December 2025)

Abstract
When do laws and policies that do not explicitly treat people differently on the basis of legally protected traits like race and sex nonetheless constitute disparate treatment on these bases? According to U.S. constitutional law, they do so when "facially neutral" laws are both enacted for impermissible reasons and also produce a discriminatory effect. To date, the first element of this claim – impermissible intention – has attracted significant attention. However, its second element – discriminatory effect – has been largely ignored. Yet it is critical that we better understand what discriminatory effect requires, as competing tests animate debates in Circuit court cases and the issue has recently been flagged by Justice Alito. This Article takes up the task. It explores the normative disagreement that underlies the controversy regarding how to assess whether discriminatory effect is present and diagnoses the genuine moral conflict that any test for discriminatory harm must navigate.

I. Introduction

As a matter of constitutional law, the United States currently recognizes two types of unlawful discrimination. First, it is unlawful to treat people differently on the basis of particular protected traits or grounds unless doing so serves a compelling interest (disparate treatment). If a law, policy, or action by an individual does not treat people differently on the basis of a protected ground, it is termed "facially neutral." Facially neutral policies are generally legally permissible, and the distinctions they draw are not scrutinized by courts. For example, if a university decides to provide an admissions preference to athletes, as compared to nonathletes, they are free to do so. But sometimes facially neutral policies give rise to legal claims. Where a facially neutral policy negatively affects a group defined by a protected trait, this fact *alone* does not give rise to constitutional concern.[1] However, when a facially neutral law has both a "discriminatory effect" and is enacted with an "invidious intent," then such a law also

[1]Washington v. Davis, 426 U.S. 229 (1976); Pers. Adm'r v. Feeney, 442 U.S. 256 (1979).

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (http://creativecommons.org/licenses/by-nc-sa/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

gives rise to strict scrutiny and is treated in the same manner as a law that explicitly treats people differently on the basis of protected traits like race and sex. We call this second form of discrimination claim "disparate treatment by proxy." It is that claim that will be our focus here.

Currently, courts understand the claim of *disparate treatment by proxy* to require a showing of two elements. Plaintiffs must prove both that the alleged discriminator had an impermissible intention and that the disputed policy had a specific effect. The content of both of these elements—that is, the matter of *which* intentions and *which* effects are legally problematic—is contested. Consider intentions first. All agree that deliberately adopting a facially neutral policy in order to harm (anyone) is a prohibited intention. What isn't clear is whether it is also prohibited to deliberately adopt a facially neutral policy in order to select people on the basis of race, sex, or some other protected trait. For example, suppose a university adopts an admissions preference for first-generation college students and does so *because* this preference is likely to increase the racial diversity of the class. In such a case, the university administrators do not intend to harm anyone, so they would not run afoul of the first understanding of prohibited intentions. They would, however, run afoul of the second understanding of prohibited intentions as the first-generation preference was adopted in order to select more students from particular racial groups. The question of whether the Equal Protection Clause is best read to prohibit the latter intention as well as the first is a critical question that the Court has not yet resolved. However, it is an issue that has already received significant attention[2] and will not be our focus here.

Instead, this Article will focus on the comparatively neglected element of *disparate treatment by proxy*: the effect prong. How is this element satisfied? While there are very few cases that address this question to date, those that do so proceed by asking whether the group allegedly discriminated against is, as a result of the policy in question, underrepresented as compared to its share of the relevant population (the population of applicants to a school, for example). Critics of this approach propose instead that courts ask whether the facially neutral policy decreases the representation of the group allegedly discriminated against, as compared to their share before the adoption of the challenged policy.

While this debate about how to satisfy the effect prong has largely flown under the radar, that is changing. The Supreme Court's recent decision in *Students for Fair Admissions v. Harvard* [*SFFA*] limited the ways in which universities can use race in the admissions process. As a result, there is increasing interest in the use of facially neutral policies to achieve the goals of student body diversity. For this reason, the claim of *disparate treatment by proxy* and its component parts (the intent prong and the effect prong) are now salient. In addition, the answers that courts provide as to how these elements are satisfied are likely to apply far beyond the university admissions contexts, as the use of facially neutral policies to achieve aims that have a racial dimension, like increasing diversity or decreasing racial disparities in various contexts, is common.

[2] *See*, e.g., Deborah Hellman, *Diversity by Facially Neutral Means*, 110 Va. L. Rev. 1901 (2024); Sonja Starr, *The Magnet School Wars and the Future of Colorblindness*, 76 Stan. L. Rev. 161 (2024); Jonathan P. Feingold, *Affirmative Action After* SFFA, 48 J. Coll. & Univ. L. 239 (2023); Brian T. Fitzpatrick, *Can Michigan Universities Use Proxies for Race After the Ban on Racial Preferences?*, 13 Mich. J. Race & L. 277, 283 (2007); Richard A. Primus, *Equal Protection and Disparate Impact: Round Three*, 117 Harv. L. Rev. 493, 496 (2003); Kim Forde-Mazrui, *The Constitutional Implications of Race-Neutral Affirmative Action*, 88 Geo. L.J. 2331, 2364–65, 2377 (2000).

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

In the dispute about how to assess whether the effect prong of the claim of *disparate treatment by proxy* is satisfied, both sides claim that their preferred approach is legally and normatively superior and that the other side's view has significant flaws. We think that both sides are correct to some extent in pointing to the virtues of their preferred approach. In what follows, we offer a fuller account than is offered in existing case law of both the benefits and burdens of each proposed test. This analysis reveals a dilemma. Briefly, the problem, we claim, is that the two accounts register different normative goods—that of affirming the forms of agency that develop in a world marked by injustice and that of making the world anew in light of that injustice—that are, in many cases, in profound conflict. This alternative understanding of the choice between the two ways of assessing the effect prong has several upshots. It, for one, avoids the simplistic and politically problematic framing that pits marginalized groups against each other in the broader struggle for equality. Second, it reveals that the law faces a starker choice that goes beyond simply selecting between these two competing ways of assessing a discriminatory effect.

Our aim is to show that each of the proposed standards of discriminatory effect, as well as the conflict between them, has deeper normative roots than it might seem at first sight. We should note at the outset that this approach means taking these battles over discrimination law at face value, as trading in reasons and guided by the moral values that are purportedly at issue. Of course, the extent to which this interpretation matches the actual motivations and intentions of the parties involved will vary by case. But the spirit of our engagement is to see what genuine moral value and normative force might lie behind legal disputes, and we take it that this approach is fruitful insofar as it successfully illuminates some features of our moral landscape that we might previously not have noticed. Thus, in taking this tack, we do not mean to endorse the broader—and in our view, dubious—view that American discrimination doctrine always presents us with a promising starting point from which to morally theorize about discrimination. It is certainly not an a priori truth that legal disputes over discrimination express or otherwise provide insight into our normative concerns about discrimination. Nevertheless, as we will argue, there are insights to be drawn from the recent struggles over *disparate treatment by proxy* and attending to them sheds light on underexplored tensions in our moral thinking about discrimination.

In order to make the discussion concrete and to illustrate the timeliness of the question we investigate, we use the recent case *Coalition for TJ v. Fairfax County School Board*[3] [*Coalition for TJ*] as an illustrative example. This case focuses on the admissions policy at one of the premier magnet schools in the country. But *Coalition for TJ* is not a lone example. Other magnet school admissions policies have also been challenged for similar reasons,[4] and the answers to the questions these cases present will have implications for the legal permissibility of using facially neutral policies to reduce racial disparities or improve racial diversity in universities and many other domains.

[3]Coal. for TJ v. Fairfax Cty. Sch. Bd., 68 F.4th 864 (4th Cir. 2023), *cert. denied*, 218 L.Ed.2d 71 (2024).

[4]*See* Christa McAuliffe Intermediate Sch. PTO, Inc. v. De Blasio, 627 F. Supp. 3d 253 (S.D.N.Y. 2022) (The plaintiffs in this case later joined as a Plaintiff in the *Chinese Americans Citizens Alliance* case, discussed *infra* notes 24–27 and accompanying text.); Bos. Parent Coal. for Acad. Excellence Corp. v. Sch. Comm. of Bos., 89 F.4th 46 (1st Cir. 2023) ("Boston Parent II"), *cert. denied* 145 S.Ct. 15 (2024); Ass'n for Educ. Fairness v. Montgomery Cty. Bd. of Educ. 88 F.4th 485 (4th Cir. 2023); Sargent v. Sch. Dist. of Phil., No. 24-3112, WL 263486 (3d Cir. Feb. 2, 2026).

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

The argument in this Article proceeds as follows. We begin in Section II with a discussion of *Coalition for TJ*, as it illustrates the issue we identify. Building on the criticisms of each manner of assessing discriminatory effect articulated in that case, Section III explains why both standards, which we term the "before/after comparison" or BAC and the "population comparison" or PC, are flawed means of assessing discriminatory harm. Given these flaws, Section IV briefly considers and rejects the possibility that the effect element of the claim of *disparate treatment by proxy* should be abandoned. Section V then turns from criticism to recovery. It articulates the genuine moral value that underlies both approaches. Unfortunately, in a world marred by injustice, these moral values are in tension, as Section VI explains. Section VII concludes.

II. Coalition for TJ v. Fairfax County School Board

In *Coalition for TJ*, a three-judge panel of the Fourth Circuit rejected the claim that the decision by the Fairfax County School Board to change the admissions procedure at the highly competitive Thomas Jefferson High School (TJ) was impermissible race discrimination in violation of equal protection. Neither the old admissions procedure, nor the new procedure adopted in 2020, used race explicitly. Indeed, the admissions files even omitted the names of applicants to ensure that the process was race-blind, in the sense of being blinded to individual applicants' racial statuses. Still, the complaint alleged that the change from one facially race-neutral policy to another was unconstitutional because the change was motivated by a desire to reduce the number of Asian American students admitted to TJ and in fact achieved that effect. On a motion for summary judgment, the lower court had held that the School Board's decision to change its policy violated equal protection. The Fourth Circuit reversed and granted the School Board's cross motion for summary judgment, finding no constitutional violation.

This case raises both the intent question and the effect question flagged above. First, the case raises the question of what intentions are constitutionally impermissible. In particular, given a fixed class size, is the intention to increase the enrollment of Black and Latino students (which the School Board acknowledged was part of its goal in changing the policy) the constitutional equivalent of an intention to "racially balance" the school or to decrease the enrollment of Asian American students? Second, even if an impermissible intent is present, how should courts assess whether the *effect* produced by the policy change establishes a constitutional violation? As we noted in the introduction, while the first question is extremely important, it has already attracted significant attention. The second question, by contrast, has so far largely been ignored.

But that soon may change. The Coalition for TJ [the Coalition] petitioned the Supreme Court for certiorari in the case, which was denied.[5] However, in a dissent from the denial of cert, Justice Alito focused exclusively on this second question, and noted his disagreement with the manner in which the Fourth Circuit assessed whether the change in policy had produced a discriminatory effect.[6] The Fourth Circuit had concluded that the Coalition could not demonstrate the new policy

[5] *Coal. for TJ*, 68 F.4th 864.
[6] *Id.*

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

"purposefully discriminates against Asian American students" unless they could show "(1) that the policy exacts a disproportionate impact on a certain racial group, and (2) that such impact is traceable to an 'invidious' discriminatory intent."[7] According to the Fourth Circuit, the correct way to analyze whether the policy produces this "disproportionate impact" (the effect prong) is to compare the percentage of Asian American students who were admitted under the new policy to the percentage who had applied to the school. And because Asian American applicants to TJ under the new policy were 48.59% of the applicant pool and 54.36% of the admitted class, the Fourth Circuit found that there was no disproportionate impact.[8] As a result, the case was without merit.

By contrast, the District Court, dissenting Circuit Judge Rushing, and Justice Alito all favored a different way of assessing whether the relevant "disproportionate impact" was present. Rather than assess how the admitted class compares to the applicant pool, these jurists favored comparing the percentage of Asian Americans in the admitted classes under the *previous* admissions policy with the percentage of Asian Americans admitted under the *new* policy. This "before/after" comparison, they claim, is what shows the relevant effect. And indeed, whereas in the five years prior to the policy change, Asian American students received 65–75% of the admissions offers, in the 2021 year (under the new policy), they received only 54%, which Judge Rushing describes as a 26% reduction in offers of admission.[9] According to this measuring stick, the policy has a clear "disproportionate impact."

The narrow legal question at hand then is this: Which comparator is the right one to use? Should the effect prong of a claim of *disparate treatment by proxy* be assessed by comparing the percentage of admitted Asian American students to the applicant pool (or some other choice of reference population) as the Circuit Court held? Call this the *population comparison* or *PC*. Alternatively, should it be assessed by comparing the percentage of Asian American students admitted under the prior policy to the percentage admitted under the new policy? Call this the *before/after comparison*, or *BAC*. These two questions lead to broader moral questions: What is the normative rationale that underlies each of these standards? Which gives a better account of a distinctively *discriminatory* effect?

To date, the PC is the test used by most courts, though only a few have addressed this issue directly.[10] The main reason that courts adopt this approach lies in a flaw they identify with the BAC. In their view, the BAC makes the prior policy the normative default, the standard against which a claim of current discrimination is

[7] *Id.* at 871, 878.

[8] *Id.* at 881.

[9] *Id.* at 901(Rushing, A., dissenting).

[10] *See*, e.g. Bos. Parent Coal. for Acad. Excellence Corp. v. Sch. Comm. of Bos., 996 F.3d 37, 46 (1st Cir. 2021) ("Boston Parent II") (holding that Plaintiffs were unlikely to succeed in challenging the magnet school admissions policy because that policy did not have a disparate impact "as compared to a random distribution of invitations" to all applicants); Lewis v. Ascension Par. Sch. Bd., 806 F.3d 344 (5th Cir. 2015) (holding that the father of two Black students, who relied on a before/after theory, failed to show that the School Board's redistricting plan had a racially discriminatory effect); Vaughns v. Bd. of Educ. of Prince George's Cnty., 574 F. Supp. 1280, 1304 (D. Md. 1983), *aff'd in part*, *rev'd in part on other grounds*, 758 F.2d 983 (4th Cir. 1985). But see Sargent, WL 263486, at 15-16 (vacating a lower court's decision after finding a discriminatory effect using a BAC analysis). Another possible exception to this consensus is *Chinese American Citizens Alliance of Greater New York v. Adams*, 116 F.4th 161 (2d Cir. 2024), which explicitly rejected the need to find either test satisfied but which implicitly relied on the BAC, as we explain *infra*.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

assessed. And if the prior policy is itself unjust, there seems to be something problematic about such an approach.[11]

We appreciate this problem with the BAC, and indeed we develop and expand on it below. But it is not dispositive of the issue, in our view, as there are powerful reasons to reject the PC as well. Suppose that members of a School Board are genuinely motivated by racial animus. The School Board thinks "we have too many Asians," "they are taking over," and accompanies those thoughts with all sorts of negative epithets about Asians that show genuine and deep animus toward Asian Americans. If the PC governs, the School Board would be free to adopt a facially neutral policy that is designed to reduce the Asian population at the school, and that does significantly reduce it, *so long as* the percentage of Asian American students at the school remains at or above their percentage in the relevant population. Both Justice Alito and the dissenting judge on the Fourth Circuit emphasize this noxious upshot of using the PC in their objections to the Fourth Circuit's approach.

Resolution of the debate about which comparator a court should use to assess a claim of *disparate treatment by proxy* depends on how one understands the nature of discriminatory harm. We should clarify, however, that this element should not be confused with the more generic legal requirement of standing, according to which the person or entity bringing the legal challenge must have suffered a harm. Any applicant to TJ who was not admitted has standing to challenge its admissions policy.[12] Nor should it be confused with the requirement that the impermissible intent must have led the relevant actor to actually do something (adopt a policy, maintain a policy, etc.). The forbidden intention cannot simply be floating around in the mind of the actor without giving rise to an *action* in the world. After all, a plaintiff must have some policy to point to that she alleges was adopted or maintained with an impermissible intent. But these general requirements of standing and action differ from the more specific requirement that the challenged policy has a particular *kind* of effect, which both the BAC and PC standards point to. These tests do not function simply to ensure that a plaintiff has standing or that the actor whose action is challenged has taken some action. Rather, as we argue below, each standard puts forth a particular conception of *discriminatory harm.* That is, the BAC and PC tests each propose an answer to the question of what kind of "effect" in the world constitutes a distinctively discriminatory effect.

III. The puzzle of discriminatory harm and the flaws of both methods of assessing it

In this section, we will probe the conception of *discriminatory harm* that underlies each of the BAC and the PC. The idea of a discriminatory harm is novel to the

[11]The Fourth Circuit puts the point this way, referring to the before/after baseline: "That approach would simply turn 'the previous status quo into an immutable quota,' thereby opening a new policy that might impact a public institution's racial demographics – even if by wholly neutral means – to constitutional attack." *Coal. for TJ*, 68 F.4th at 879 (quoting Brief of Appellant at 25).

[12]Standing doctrine, interestingly, does not require that the harm suffered be impermissible under the legal doctrine at issue. In other words, a plaintiff can have standing in cases in which their legal claim fails on the merits. *See* Gerber v. Herskovitz, 14 F4th 500, 505 (6th Cir. 2021). ("The standing inquiry is not a merits inquiry. A merits defect deprives a court of subject matter jurisdiction only if the claim is utterly frivolous." (citing Steel Co. v. Citizens for a Better Env't, 523 U.S. 83, 89 (1998))).

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

particular kind of discrimination under study—*disparate treatment by proxy*—and the normative question at hand is what exactly it is. When does some outcome or effect amount to a distinctively *discriminatorily* harmful outcome or effect? As noted above, this question should be distinguished from the more generic question of whether adopting some policy effects any harm at all, broadly construed. In TJ, that more generic standing requirement may be satisfied so long as the admissions policy at issue denies *someone* entrance to the school. The question, though, is whether the harm of being denied entry into one's school of choice is legally cognizable as a discriminatory harm. To answer this requires an account of discriminatory harm.

To better understand the nature of this inquiry into discriminatory harm, it might help to contrast it with other cases where harm is at issue. In many cases, it is obvious that a given state of affairs constitutes a harm and so we neglect to draw out explicitly the basis on which the state of affairs is a harmful one. For instance, when I break your arm, I effect a state of affairs in which your arm is now broken. We can immediately recognize this to be a harmful state of affairs, without working through some underlying theory of harm to explain why. This is not to say that there is no further explanation why having one's arm broken constitutes a harm; there certainly is. But in cases like this, we can readily identify that the state is harmful without filling in a particular account on which it is. By contrast, in the cases at hand, where the effect of some policy is to simply admit some set of students and deny others, it is not only not immediately clear whether the resulting state of affairs constitutes a discriminatory harm, but it is not even clear how one should go about assessing this. Hence we have in TJ a debate between two different empirical tests, which emerge out of two different substantive conceptions of what makes some outcome a case of "discriminatory harm." Each test characterizes the effect that a school admissions policy has on the demographics of its student population in a distinctive way, and thereby proposes that *this* way of conceptualizing the policy's effect brings out what is potentially discriminatory about it. According to the BAC, determining whether some policy produces a discriminatory harm requires comparison of its resulting demographic profile with that of a previous policy. For the PC, the comparator should instead be the demographics of some catchment area. Scrutinizing the BAC and PC standards thus means asking why each standard's approach to characterizing some effect—say, as a *decrease*, or as *underrepresentation*—should be what draws out what is distinctively morally significant about it (vis-à-vis discrimination). Unfortunately, the conception of discriminatory harm that underlies each standard is normatively flawed as we elaborate below, beginning with the BAC.

To understand the normative vision of *discriminatory harm* that underlies the BAC, a good place to begin is with Justice Alito's dissent from the denial of certiorari in the TJ case. He writes:

> Suppose that white parents in a school district where 85 percent of the students are white and 15 percent are black complain because 10 of the 12 (83%) of the public high school basketball team are black. Suppose that the principal emails the coach and says: "You have too many black players. You need to replace some of them with white players." And suppose the coach emails back: "OK. That will hurt the team, but if you insist, I'll do it." The coach then takes five of his black players aside and kicks them off the team for some contrived—but facially neutral—reason. For instance, as cover, he might institute a policy that reserves a set number of spots on the roster for each of the middle schools who

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

> feed to the high school. According to the reasoning of the Fourth Circuit majority, this action would not violate equal protection because the percentage of black players left on the team (approximately 42 percent) would exceed the percentage of black students in the school.[13]

If you think this example describes a discriminatory harm, as Justice Alito does, then perhaps the BAC captures something important. But what?

The answer isn't simply that the policy was enacted with a legally problematic intention (however defined). Rather, there must also exist a distinctively discriminatory *harm*. The BAC takes the harm to inhere in the fact that there is a *decrease* in the number (or percentage) of Black players on the basketball team. Now, it is certainly true that the policy results in a decrease in the number of Black players on the team. But the question at hand is *why* the mere fact of a decrease should make it the case that the resulting state of affairs realizes a "discriminatory harm"? What we need is an argument for *why* this temporal comparison between the present composition of the team and the preceding composition grounds a claim of discriminatory harm. Otherwise, claiming that the policy has a discriminatory effect because it produces a decrease relative to the prior policy amounts to assuming rather than justifying that the BAC provides the right baseline for determining what constitutes discriminatory harm.

The above discussion thus imparts a methodological lesson. To avoid accidentally presupposing a substantive account of discriminatory harm in examining the effect that a policy has, we should try to characterize the outcome in as neutral terms as possible. So rather than describing the policy as "decreasing" or "under-representing" some group—terms which implicitly refer to a particular substantive choice of baseline—we should instead characterize the policy as *resulting in X racial breakdown*. Recasting the effect in this way helpfully brings out the ambiguity in how such states of affairs should be assessed: It is simply unclear whether a basketball team with 5 Black players out of a total 12 players *really is* one in which Black students are discriminatorily harmed.

Returning to the BAC in particular: notice that in setting forth a *temporal* or *historical comparative* account of harm, its animating normative picture of discrimination is one that is concerned with the change in a racial group's representation over time.[14] In particular, in characterizing policy outcomes in terms of how they relate to past policy outcomes, the BAC both privileges the status quo and treats the status quo as a normative baseline for determining whether and when some state of affairs constitutes a discriminatory harm. Status quo entrenchment is a problematic feature of institutional design, and especially so in domains where concerns about discrimination arise. This aspect of the BAC would thus seem to pose serious risks and pitfalls for its account of discriminatory harm. To better see what these are, it will help to break up consideration of the BAC into two types of cases. First, there are cases in which the state of affairs before is the same as the state of affairs after the action at issue. Second, there are cases in which the state of affairs before differs from the state of affairs after the relevant action. We discuss each scenario below, with the aim of

[13]Coal. for TJ v. Fairfax Cnty. Sch. Bd., No. 23-170, slip op. at 9 (U.S. Feb. 20, 2024) (Alito, J., dissenting from the denial of certiorari).

[14]For defenses of the temporal comparative account of harm, *see* Stephen R. Perry, *Harm, History, and Counterfactuals*, 40 San Diego L. Rev. 1283–1313 (2003).

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

highlighting how the BAC standard operates to preserve the status quo and the key shortcomings that follow from this conception of discriminatory harm.

First consider a case in which there is no change in outcome. Since the BAC asks that courts compare how a group fared *before* some poorly motivated action occurred with how the group fares *after* that action, where there is no difference between how the group fared before and after the action at issue, there can be no violation. For this reason, the BAC enacts a status quo bias, in both a descriptive and normative sense.

In particular, where there is no change in the world, there can be no constitutional violation because the BAC requires that a group's share of the admissions spots (or whatever else is at issue) *diminish.* As a result, this test cannot recognize poorly motivated decisions that yield no change. In particular, poorly motivated decisions to maintain policies are good examples of cases that are especially likely to evade scrutiny under this test, because such decisions are likely to yield no change in results. To illustrate, consider the following case.

Scrapped Plans

> A school board votes to alter its longstanding policy on entrance exams at the recommendation of an independent committee formed to evaluate the school's admissions policies. But the plan is scrapped at the last minute upon the board's finding that the new policy is projected to lead to more Black and Latino students being admitted. So in the end, the board maintains the prevailing policy. And, as a result, the demographics of the admitted class remains approximately the same.

In declining to adopt the new policy on the grounds that it is likely to increase the admission of Black and Latino students, the school board in Scrapped Plans has, in our view, made a decision that is morally troubling. Moreover, and perhaps more importantly, Scrapped Plans seems as morally troubling as the decision to adopt a *new* policy for the same reasons. Because new policies are more likely to result in changes in who is admitted than decisions to maintain policies, the BAC is significantly more likely to find a violation in the second case than in the case of Scrapped Plans. Yet this asymmetry in how the BAC treats the two cases seems unjustified.

The second manner in which the BAC reinforces the status quo is more distinctively normative. When the BAC serves to test for some legal-normative violation, the "before" baseline establishes the normative benchmark to which all new policies are compared and in light of which they must be justified. This framing has two consequences. First, any future policy can never be considered on its own merits and must instead, always, be characterized as a *departure* from a previous policy and judged on those terms. Second, since the prevailing policy constitutes the yardstick of constitutional violation, the current policy cannot itself be scrutinized in the same way.[15] That is, whether the prevailing state of affairs is considered independently bad or good is immaterial to the BAC. Instead, prevailing policies earn their keep simply by having already been instated. In Scrapped Plans, for instance, the BAC would

[15]This contrasts notably with the PC, which applies equally to both the "before" and the "after" policies. The demographic profile of admitted students can always be measured against some reference population and the result of such a test and its legal consequences will be wholly insensitive to any facts about other admissions policies.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

focus on the board's decision to change its longstanding policy, while this policy itself receives no such scrutiny. So, the BAC in effect freezes the school's policy at an arbitrary moment in time (the time of the court's adoption of this standard, or perhaps the policy immediately preceding the policy that is being challenged) and proclaims that policy to be de facto justified. On this way of looking at it, the BAC does not just risk or embed a bias toward the status quo. Rather, to the extent that it effectively defines nondiscrimination by reference to the reasons behind and the effects of the prevailing policy—without ever bringing independent scrutiny to those reasons and effects themselves—it just *is* a status quo preserving standard.[16]

Next consider the situation in which the outcome does change over time. When the outcome in the world changes after a poorly motivated decision (likely because a new policy is adopted), the BAC may also entrench the status quo. To illustrate, consider the following case:

Policy Change

> A School Board decides to change the admissions policy for a public magnet high school. Under the old policy, admission to the school was determined by results on a standardized test. Under the new policy, the top 1.5% of the eighth-grade class at each middle school in the catchment area is admitted to the magnet high school. This change is adopted in order to increase the percentage of Black and Latino students at the school.

The first thing to note about contexts like this one in which admissions are competitive and spots are limited is that *any* policy change is likely to change the school's demographics. And any change in demographics results in some group's percentage of the admitted students increasing and some group's percentage of the admitted students decreasing. In other words, in virtually all cases of policy changes, there will be some group whose percentage of the school's population has decreased in comparison to what it was before. The BAC will thus nearly always set the stage for a complaint by some group to any policy change. If decisions that maintain the status quo are likely not open to criticism and policy changes are likely to always risk criticism, the BAC seems to doubly entrench the status quo. And in a world marked by past and present injustice, to entrench the status quo just is to entrench many of these injustices.

Of course, a decrease in a group's percentage among the admitted students does not, by itself, give rise to a legal claim of *disparate treatment by proxy*—even if we grant that the BAC presents the relevant standard for determining a discriminatory effect. Nonetheless, it is important to note that the BAC ensures that the effect element of the claim will be satisfied for some group in nearly every case (the only exception being the rare policy change that does not alter the demographics of the admitted student body in any way).

[16]This line of thought shares much with critiques of temporal accounts of harms that point out that they are unable to recognize cases of congenital harm, in which a person is in a state of being badly off from birth. Just as it seems unintuitive to say that a person who was born into conditions of impoverishment does not suffer harm at the hand of those conditions, it seems unintuitive to say that a group which has suffered discrimination since a school's founding does not now suffer a discriminatory harm since they are not any worse off than they've ever been in the past.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

Next consider the PC. Whereas the BAC expresses a temporal conception of discriminatory harm, which favors stasis over time, the PC proposes that an effect be evaluated by reference to a particular normative ideal set by the demographics of some relevant population, say, the demographics of a school's catchment area. On this view, a policy should select a population that roughly reflects the demographic composition of this reference population, and departure from such proportional representation constitutes a discriminatory harm for those groups that are underrepresented.

This measure of discriminatory harm, like that adopted by the BAC, requires a justification as it is not obvious why the fact that a group's allocation of spots in a school (for example) is less than its representation in the catchment *is* a discriminatory harm. The fact that a particular pattern exists between these two measures is not meaningful on its own as either a harm or, specifically, a discriminatory harm. So, the first objection one might have to the PC is that it lacks a normative grounding as it does not explain why we should object to the failure of the school's demographic profile to match the population from which it draws.

One could, of course, flesh out this account by providing such an explanation. In our view, the most plausible account of why a departure from proportional representation would constitute a discriminatory harm is the supposition that any such departures have their roots in prior discrimination. For example, the thought is that were it not for racial discrimination in the past, the racial breakdown of the student population at the magnet school *would* match the racial demographics of the catchment area. And thus any departure from this discrimination-free population baseline constitutes a discriminatory harm. If some thought along these lines underlies the PC (an idea we develop in Section V), the PC is thereby committed to something further: that a policy fares *better* in terms of discriminatory effect, the closer it gets towards this ideal state.

Let's grant, for the sake of argument (for now), that mirroring the demographics of a given reference population is in fact ideal. Still, understood in such broad terms, the claim that descriptive improvements toward this ideal and normative improvements move in lockstep is unlikely. To return to Justice Alito's hypothetical, it certainly does not seem as though adopting a policy that leads to an increase in the representation of white students on the basketball team would make for a better state of affairs than maintaining the current basketball try-out regime, simply because it is descriptively closer to the racial composition of basketball teams in an ideal world. But this is implicitly what the PC approach assumes. It assumes that because a given standard is the ideal state, those states of affairs that depart less from this state are preferable to those that depart from it more. In fact, this is the reason why the PC judges that in TJ, a policy change that reduces the percentage of Asian American students without bringing them below the population baseline is a change that does not make things worse (and perhaps even makes things better). But this conclusion strikes us as too quick. In the real world, in which prior injustice has occurred, it is not obvious that this is an improvement. In fact, by the lights of the PC, it would even be an improvement if the Asian student population decreased in favor of more white students, holding fixed Black and Latino representation, since that change would also bring the profile of admitted students more in line with the demographics of the applicant pool.

David Estlund calls this error the "fallacy of approximation."[17] This fallacy rests on the thought that "if you have some ingredients of a good condition, but you will not

[17]David M. Estlund, *The Fallacy of Approximation*, *in* Utopophobia: On the Limits (if Any) of Political Philosophy 271, 271(2020). Princeton: Princeton University Press.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

have all of the ingredients, having more of them would be better."[18] However, while this is sometimes the case, it is not always the case. Estlund illustrates the fallacy with the example of a person who is prescribed three pills for a medical condition but cannot get access to all three. If she can only take pills 1 and 2 but not pill 3, it is not clear that taking pills 1 and 2 is better (for her health, say) than taking only one pill or none at all. We would need to know more about the pills, how they work, and the interaction among them to make that assessment.

The "negative implication" of the fallacy of approximation is perhaps even more relevant for the scenario we are investigating, according to which "understanding some high standard is not useful for understanding value rankings where the standard is not fully met."[19] If we bring that insight to bear in the case we are considering, we might say the following. The fact (supposing that it is a fact for the sake of argument) that in the ideal world, in which no prior discrimination had occurred, the racial breakdown of the population of admitted students would roughly mirror the racial breakdown of the geographic area the magnet school draws from does not in itself help us to discern which nonideal demographic breakdowns are better or worse than others. For this reason, even if we grant that the PC correctly expresses the ideal, it does not tell us whether any particular policy change is a normative improvement.

IV. Should the discriminatory effect element be abandoned?

The prior section highlighted the ways in which each of the conceptions of discriminatory harm present in the case law has significant problems. The BAC treats any departure from the present state of affairs as presumptively problematic and thereby enacts a form of status quo bias that is unjustified. The PC, by contrast, rejects stasis as a normative ideal in favor of an imagined ideal standard. But it too is flawed because it mistakenly assumes that states of affairs that are closer to this ideal are necessarily better than those that are further away. Given these significant problems with the two current approaches for assessing discriminatory harm, perhaps this element should be abandoned. If it were, the claim of *disparate treatment by proxy* would only require a showing of impermissible intent and no showing of discriminatory harm at all. (Though, as we noted earlier, plaintiffs would still need to have standing to bring such a claim.)

While this *intent only* version of the *disparate treatment by proxy* claim is surely possible and there is some language in the cases that supports it, we should note that all courts considering the claim say that a plaintiff must show both impermissible intention and some sort of impact.[20] For example, in *Coalition for TJ* itself, the Circuit

[18] *Id.* at 273.

[19] *Id.* at 279.

[20] Lewis v. Ascension Par. Sch. Bd., 806 F.3d 344, 358–59 (5th Cir. 2015) (holding that because a discriminatory effect was not shown, there was no need to assess whether discriminatory purpose is present and insisting that to "subject a facially race neutral government action to strict scrutiny, the plaintiff must establish both discriminatory intent and a disproportionate adverse effect upon the targeted group"); Bos. Parent Coal. for Acad. Excellence Corp. v. Sch. Comm. for Bos., 89 F.4th 46, 59 (1st Cir. 2023) ("[T]o successfully challenge the use of a facially neutral, and otherwise bona fide, selection criterion, the [plaintiff] must prove both improper intent and disparate impact."); Hayden v. Cnty. of Nassau, 180 F.3d 42, 50 (2d Cir. 1999) ("[A]ppellants must sufficiently allege that the Nassau County defendants harbored a discriminatory

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

Court first assessed whether the requisite impact was present and, having found that it was not, noted that there was no need to go on to consider whether there was sufficient evidence of improper purpose (though it did so nonetheless).[21] If an effect were unnecessary, then failure to satisfy the effect prong would not suffice to invalidate the claim. The fact that the Fourth Circuit, and other courts, require some effect demonstrates that the *intent only* version of the claim does not fit existing case law.[22]

That said, the case law assessing claims of this form is continuing to evolve.[23] Indeed, the United States Court of Appeals for the Second Circuit recently decided a case dealing with admissions policies at the New York City magnet schools that might be read to hold that a plaintiff need only show an intent to decrease the Asian population at the schools and the denial of admissions to some Asian students.[24] In *Chinese American Citizens Alliance of Greater New York v. Adams [CACAGNY]*,[25] the posture of the case requires the court to assume that impermissible intent is present and so the sole issue on appeal was whether the plaintiffs have established the requisite "discriminatory effect." The intention that is presumed is that the city intended to decrease the percentage (or number) of Asian students in the New York City magnet schools. Because the new policy led neither to a decrease in the Asian population at the schools (and thus did not satisfy the BAC), nor was the Asian population in the magnet schools lower than the percentage of Asian students in the school-age population eligible to apply to the magnet schools (and thus neither did it satisfy the PC), the trial court held that the plaintiffs had not stated an equal protection claim. On appeal, the Second Circuit reversed, finding that discriminatory effect should also be evaluated on an individual basis rather than only on the group or aggregate basis instantiated in either the BAC or PC. Using this approach, the court found that the requisite discriminatory effect was shown by the fact that some Asian applicants who would have been admitted under the prior policy were rejected under the new policy.

Here's how the court describes its holding: "when a discriminatory intent in enacting a facially neutral law or policy is shown, the requisite proof of discriminatory effect is not limited to an aggregate disparate harm, but rather can be satisfied by proof of an adverse harm to one or more individuals based on their race that has

intent against them and that the entrance examination disproportionately impacted them."); Ass'n for Educ. Fairness v. Montgomery Cnty. Bd. of Educ., 617 F. Supp. 3d 358, 368 (D. Md. 2022), motion for relief from judgment denied, No. 8:20-CV-02540-PX, 2022 WL 17740402 (D. Md. Dec. 16, 2022) (finding that the lack of plausible showing of disparate impact and discriminatory intent provides alternative bases to reject the plaintiff's claim).

[21]Coal. for TJ v. Fairfax Cty. Sch. Bd., 68 F.4th 864, 882 (4th Cir. 2023).

[22]Indeed, in *Palmer v Thompson*, 403 U.S. 217 (1971), the Supreme Court specifically held that when assessing whether facially neutral governmental actions (in that case, closing the town's swimming pools) violate the constitutional guarantee of equal protection, it is necessary to show discriminatory effect in addition to invidious intent. Because pool closures meant no one could swim at city pools, the Court found that there was no discriminatory effect and so no equal protection violation.

[23]Moreover, *Palmer v. Thompson*, described above, was a long time ago and dealt with a policy that did not treat people differently on the basis of a neutral trait as the cases we are examining do, rather it treated everyone the same. For that reason, an argument could be made to distinguish that case from the ones under consideration.

[24]Chinese Am. Citizens All. of Greater New York v. Adams, 116 F.4th 161 (2d Cir. 2024).

[25]*Id.*

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

resulted from that discriminatory law or policy."[26] But to the extent that this is offered as a conception of discriminatory harm and not merely as a reiteration of the standing requirement, this statement is puzzling and, ultimately, circular. While an Asian student plaintiff has indeed been denied admission, what makes this effect "based on their race" and thus a *discriminatory* effect rather than merely the sort of harm that is necessary for standing? After all, the policy at issue does not employ race. Perhaps the claim rests on the fact that the student is denied admissions *and* is Asian. But there are many students of different races who were denied admissions under the new policy (and perhaps would have been admitted under the old policy). Why exactly have these other students not also stated a claim for a violation of equal protection? After all, the policy under which they were denied admission was motivated by an impermissible intention and they suffered harm. So what exactly makes the harm "discriminatory" in the case of an Asian student but not "discriminatory" in the case of a student of another race? Granting that the facially neutral policy was adopted with the invidious intention of harming Asians, how does the content of that intention make a difference to the nature of the harm that is suffered by different individuals on account of the same policy? Setting aside matters of blame or culpability, it would appear that the *harm* of being denied admission is the same, regardless of who it is that is being denied admission. The court in *CACAGNY* provides no answer for why it is that only Asian students suffer the discriminatory harm. For this reason, this case would seem to suggest that the *discriminatory effect* element is unnecessary.

But when we dig deeper into the reasoning of this case, we see that the court implicitly relies on the BAC baseline after all. Notice that the court in *CACAGNY* does not treat all Asian students who were denied admissions under the new policy as having suffered a harm "based on their race." Rather, it is only those who would have been admitted under the prior policy who qualify as having suffered the requisite discriminatory harm.[27] This restriction on *which* Asian students may bear proper witness to the policy's discriminatory effect shows that the court in *CACAGNY* implicitly relies on a version of the BAC in assessing what constitutes discriminatory harm, albeit in an individualized rather than an aggregate population-wide form. Perhaps the court does so because it realizes that the BAC fills in a conception of discriminatory harm that is otherwise missing in its analysis.

As this discussion illustrates, existing case law maintains the importance of showing *discriminatory harm*—though the question of what such a harm *is* has remained persistently ambiguous. In addition, while an intent-only version of the claim of *disparate treatment by proxy* remains a theoretical possibility, we are not yet ready to abandon the necessity of showing discriminatory harm. In part, this judgment rests on the very real values on which it rests, as we explain in the next Section.

V. Genuine discriminatory harm as captured by both the BAC and the PC

While both the BAC and the PC articulate flawed conceptions of discriminatory harm, as we explored in Section III, there is a kernel of truth that each captures as well.

[26] *Id.* at 173.

[27] For example, the court focuses precisely on such a student when explaining that given the posture of the case in which they must presume discriminatory intent, "we must presume that the Asian-American students who have been rendered ineligible for the Discovery Program by virtue of the middle school they attended – despite otherwise having qualified for admission on an individual basis under the previous criteria – were excluded from those designated SHS seats precisely because of their race." *Id.* at 174.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

In this Section, we explore what that kernel is. In the next Section, we will explain the tension between these two visions and how that tension presents a very real dilemma for how best to understand the concept of discriminatory harm.

In Section III, we explained that the BAC privileges the status quo. There, we focused on the normatively troubling features of status quo bias: in particular, its tendency towards the entrenchment of existing injustices. But be that as it may, this does not mean that status quo bias is wholly unjustified, as the stability of the institutions that govern access to important opportunities indeed has moral value. In particular, philosophers have long noted the pervasive significance of *reliance* in our lives. Reliance lies at the heart of rational agency itself: reliance, for example, that that which typically serves as a means to an end, will this time, when I take it, also serve as a means to my desired end. This in turn means that reliance has special normative import in a liberal society. The exercises of agency characteristic of a liberal society require individuals to have more than just the mere latitude to author and pursue their own lives. Agent deliberation, action, and planning must meet up with a broader causal environment that is sufficiently stable and predictable in order for individuals to be able to successfully author their lives. As a general matter, then, the normative significance of reliance draws from the normative significance of supporting agency itself.

Thus, from the perspective of safeguarding agents' abilities to plan their lives in rational pursuit of key institutional goods, there is some disvalue in making changes to policies because it weakens reliance—even if the change would be more rational, fairer, or even all things considered an improvement. Of course, even while this interest in stability is real, it must be weighed against countervailing interests in modifying and improving the processes by which important opportunities are allocated. This is especially pressing when current institutional arrangements are unfair or work to compound prior injustice. So, there is a real tension here that we will return to and further develop in Section VI.

Still, in some cases, reliance interests are not only of general value to all in a liberal society but failing to ensure them can even present a distinctively *discriminatory harm*. Consider the facts in *Coalition for TJ* itself. At issue in this case is the policy's effect on Asian representation within the new class of admitted students. It is notable that in our world, Asian students have historically been (and in many cases, still are) affected by discrimination in schools. For example, within the context of competitive school admissions, subjective assessments of leadership skills or personality have been applied to Asian students, in the past and in myriad settings, in a discriminatory manner. Unconscious reliance on familiar stereotypes of Asian students, e.g., passive or dull, affects Asian students' abilities to perform well on these dimensions. Suppose, now, that discriminatory harms like these in part explain Asian students' investment in admissions criteria such as standardized test scores and accordingly their relatively better performance on these measures compared with those where they are unfairly disadvantaged. If so, then the policy change away from test-based assessments would disproportionately harm Asian students in considerable part because of their past experiences with discrimination. This would seem to make the effect of the policy change on Asians discriminatory indeed. Furthermore, the "overrepresentation" of Asian students relative to their population share in a context in which test-based assessments predominate (like TJ admissions before the change) is importantly *not* an overrepresentation due to a set of prior unfair advantages that the group had.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

To support our view that, to the extent that BAC captures genuine discriminatory harm, it does so when and because it protects the reliance interests of individuals who have navigated injustice, consider a modified version of the TJ case, which we call TJ*.

> TJ*: Suppose the School Board charged with determining the admissions policy at TJ* is motivated to reduce the percentage of white students at TJ. In order to do so, they abandon the prior admissions policy according to which students were selected on the basis of a standardized test and replace it with a policy according to which the top 1.5% of the eighth-grade class at each feeder middle school is admitted. As a result, the percentage of white students declines, however they are still overrepresented among the admitted students at TJ*.

TJ and TJ* would be judged similarly under the BAC standard, yet there seems to be a normatively important difference between them. Most notably, in TJ*, it is plausible to think that the status quo of overrepresentation of white students has come about because of unfair advantages these students and their families had access to. For this reason, a decrease in their representation may be a step in the right direction, because it lessens their overrepresentation as compared to what would have occurred absent prior racial discrimination (an intuition that the PC seems to capture).[28] The real TJ case, by contrast, wherein it is Asian students who are "overrepresented," feels different *because* it seems that prior and current discrimination against Asian students could plausibly explain why Asian students do especially well on standardized tests and are thus overrepresented among the accepted students under the old test-based policy (an intuition that the PC failed to register, as we explained earlier). This contrast between the real TJ case and TJ* suggests that the BAC's presumption in favor of the status quo only tracks a genuinely discriminatory effect when the group that is disadvantaged by some policy is disadvantaged today in part because of discrimination in the past.[29]

This intuition draws further support from the fact that the groups challenging policy changes under a *disparate treatment by proxy* theory do so on behalf of plaintiffs, like the plaintiffs in *Coalition for TJ*, who have themselves suffered prior discrimination. In fact, the hypothetical Justice Alito constructs to illustrate the genuine discriminatory harm that BAC registers, and PC misses, offers an example with a remarkably similar structure. In particular, it seems to us that Black students' representation on school basketball teams is not something that is entirely independent of their marginalization in other areas of school. Rather, there is very plausibly a link between the two. Perhaps opportunities to play sports were

[28]For a subgroup of White families, the success of their children may not be tied to unfair advantage, however. Families who are poor or otherwise disadvantaged would fall into this category. Where their poverty is not the result of discrimination, they would suffer a harm that has moral weight if the policies on which they rely are changed. We would not consider this to be a *discriminatory harm*, however.

[29]Contrary to our view, one might think that the purpose of the BAC is to demonstrate *causation*, in particular that the improper purpose caused the discriminatory effect. Even if the BAC did provide evidence of causation in this way, one would still need to demonstrate a discriminatory effect. And, as we explained earlier, the fact that a particular demographic profile of students is admitted cannot by itself constitute a discriminatory harm. One can only register such a demographic profile as a "decrease" in the population of some group if one *assumes* rather than defends that the BAC constitutes the yardstick for discriminatory harm. In this section, we are trying to unearth what that conception of discriminatory harm might be. Our thanks to an anonymous reviewer for raising this point.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

comparatively more open to Black young people than were opportunities to excel in other ways, or perhaps common stereotypes about the abilities and deficits of Black students worked to steer Black students into basketball rather than science, for example. If this were the case and the present-day representation of Black players on the basketball team could in part be traced back to past discrimination practices against Black students, then it *does* seem troubling to kick Black players off the team today in the name of nondiscrimination. After all, the reason Black students as a group are even in this position to be disadvantaged by the policy change in the first place is because of past discrimination.

Justice Alito is right to note that there is something indeed discriminatory about a policy that reduces a group's representation in the latter case, since the *reason why* the policy has that effect on the group is likely because of a history of discrimination against this group. When "overrepresentation" is the result of a group's *navigating past discrimination* (as might be the case with Asian students in the actual TJ case and Black students in Justice Alito's basketball example), policies that disadvantage that group on that basis today are naturally thought of as having an effect today that is a legacy of this past discrimination and so is itself discriminatory. This is the conception of *discriminatory harm* that the BAC appropriately captures.

At this point, we should make clear that we do not here mean to be positing that these causal stories about navigating past discrimination are in fact what explains the "overrepresentation" of Asian students at TJ or the hypothetical "overrepresentation" of Black students on the basketball team in Justice Alito's example. Ultimately, validating whether there is such a link is a matter of historical and social-scientific rather than normative and theoretical inquiry. Nor do we mean to suggest that concerns to safeguard reliance and attend to how groups have navigated past injustice are what in fact motivate various actors to defend the BAC. Rather, our aim has been to put forth plausible explanations for the divergence in our moral intuitions about these cases—the real TJ and Justice Alito's basketball hypothetical, on the one hand, and TJ*, on the other—and connect them to an account of why the BAC in different cases does and does not seem to latch onto something significant vis-à-vis discrimination. These explanations center on the after-effects of past discrimination on how groups navigate the social world today. The thought is that when past discrimination against a particular social group explains why it is disadvantaged by a given policy change today, these effects on the group fit within the normative category of "discrimination."

Thus, while normative discussions of the effects of preserving the status quo tend to focus on the preservation of unfair advantage, there are also circumstances under which there are good reasons to support a presumption in favor of such stability. Cases like Justice Alito's basketball hypothetical bring out a side of the prevailing state of affairs with which we have more positive associations: It also reflects how groups have managed to navigate past discrimination and the forms of agency that they have developed in response to injustice. From this vantage, the status quo contains sites of opportunity even for those social groups who are and historically have been disadvantaged. Insofar as policy changes threaten to alter disadvantaged groups' access to these opportunities, it seems to us that they should indeed face a burden.

In sum, the BAC could rest on two broad kinds of harms that track genuine concerns of discrimination. First, in a society with a history of injustice, there will be certain social resources or avenues to opportunity that disadvantaged groups have

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

come to rely on in responding to past discrimination. In these cases, policy changes which disproportionately disadvantage the group by altering access to these goods trace back to this past discrimination because the reason they became especially important to the group is the unjust ways that other avenues toward opportunity have been shut off to them. In these cases, the harm today is quite literally an effect of past discrimination and so seems to itself be a discriminatory effect. Second, seeing the status quo as embodying not just injustice but also adaptive responses to injustice and successful navigation of unjust social conditions allows us to see how policy changes may further disadvantage already disadvantaged groups. It does so by undercutting both their access to important social goods as well as their agential powers in acting and planning around key social institutions. If the moral wrong of discrimination is tied to the entrenchment of disadvantage, then these effects would straightforwardly constitute discriminatory harms.

Having unearthed what we see as the genuine moral intuition that underlies the BAC, we turn now to the PC, and, in similar fashion, attempt to recover the kernel of moral value on which it rests. Rather than compare the state of affairs after the policy change to the state of affairs before the policy change—as the BAC did—the PC compares the state of affairs after a policy change to a particular normative ideal, one in which each group is represented in rough proportion to its percentage of the relevant population. By focusing on that comparison, we can see what normative vision undergirds the PC.

In our view, the most promising normative rationale for adopting the PC is that it instantiates a view of what the world would have been like had discrimination not marred the past. Absent such discrimination, the PC seems to assume, admission to competitive high schools like TJ would not correlate with race. Rather, each racial group would be likely to be represented among the student body in proportion to their representation in the applicant pool or in the catchment area. Absent prior discrimination, kids from different racial groups would not differ from one another in terms of aptitude, ability, or interest. And so, the racial composition of the student body of an elite magnet school would likely mirror the racial composition of the geographic area from which the students come. The reason to measure a policy change with reference to the state of affairs that would have occurred absent discrimination is so that a court may assess whether a disputed policy change, even when animated by a problematic motivation, nonetheless moves us closer or further away from this idealized state of affairs. When the change moves us closer to this state, even if poorly motivated, no harm (moral or constitutional) occurs. Or so the thought might go.

It is important to note that this account of PC's normative rationale does not rest on the view that proportional representation *as such* is what fairness requires. Nor does it depend on what "really" would have happened in such an imagined counterfactual world. Rather, it rests on a sort of *legal fiction* about what the world would have been like absent discrimination. Hence, an objection claiming that even absent discrimination, there would be differences among racial and ethnic groups along multiple dimensions, perhaps due to cultural differences among different groups, erroneously takes the counterfactual claim in the vein of an empirical claim. The point instead is that in the absence of knowledge about what would have been the case in such a counterfactual world—whether, for instance, cultural differences would or would not create such differences—the law presumes that since people from different racial groups do not differ in ability or other meaningful ways, any observed

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

differences must result from unjust social practices.[30] By calling this presumption a "legal fiction," we do not mean to suggest it is false. Rather, its status as legal fiction rests on the fact that while its underlying truth is unknowable or unprovable, the law adopts it as a lodestar for normative and expressive reasons.[31]

To be sure, the PC is only a *partial* idealization of what the world would be like absent racial discrimination because the composition of the catchment area of the school itself would likely be very different from what it currently is absent prior race discrimination. The PC is not animated by a view about what the world as a whole would be like absent racial discrimination. Rather, it instantiates an epistemic commitment that absent prior race discrimination, there would be no differences in talents and abilities between racial groups and, for that reason, the population of admitted students would roughly match the racial composition of the catchment area.[32]

The kernel of normative value that underlies the PC is a commitment to a vision of racial equality in which racial differences make no difference in the aggregate to the abilities, opportunities, or achievement of students. It is a vision of a world freed from the after-effects of racial discrimination and thus, in a substantial way, freed from the significance of race in terms of station or status. This vision is borne out in a racially egalitarian distribution of key benefits and burdens, with the thought being that part of having achieved a racially just society is to have overcome the racially patterned outcomes that characterize our current social order.

VI. An unavoidable tension

We have seen in the preceding two sections how each of the BAC and PC standards expresses a view of the kind of effect that constitutes a distinctively discriminatory harm. In considering deviation from the status quo as the hallmark of discriminatory harm, the BAC risks carrying forward a history of group disadvantage and marginalization. But it also potentially pays heed to the efforts of individuals and groups at navigating that discrimination, and it suggests that there should be a presumption against policy changes which undermine those efforts. The BAC's conception of discriminatory harm suggests, reasonably in our view, that actions that disadvantage a group in a way that traces back to the group's encounters with discrimination in the past are themselves discriminatory. Furthermore, in attending to how groups adapt to existing social conditions, the BAC takes note of cases wherein policy changes would deepen the distributive harms that a disadvantaged group suffers or present a

[30]*See* Frederick Schauer, *Legal Fictions Revisited*, *in* Legal Fictions in Theory and Practice (Maksymilian Del Mar & Willian Twining, eds. 2015). Cham: Springer. (describing legal fictions as functioning as presuppositions, presumptions, and prevarications).

[31]*See* Brown v. North Carolina, 479 U.S. 940,941 (1986) (O'Connor, J., concurring in the denial of certiorari) (characterizing the Court's decision in *Batson v. Kentucky*, which rejected the use of race in peremptory challenges, as resting on "this nation's profound commitment to the idea of racial equality, *a commitment that refuses to permit the State to act on the premise that racial differences matter*" and describing that commitment as a "special rule of relevance, a statement about what this Nation stands for, rather than a statement of fact" (emphasis added)).

[32]*See* Deborah Hellman, *The Epistemic Commitments of Nondiscrimination*, in Oxford Studies in the Philosophy of Law, (Leslie Green, Brian Leiter & John Gardner, eds., 2022). Oxford: Oxford University Press.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

distinctively burdensome obstacle to a disadvantaged group's ability to rationally pursue life plans. These too raise genuine worries of discrimination. The PC, on the other hand, sets forth a view of discriminatory harm that compares the effect of a given policy to a normative standard that is drawn from a society in which, without group stratification, it is imagined that the demographic profile of recipients of scarce competitive goods would reflect the demographics of some broader population. If this is the pattern of outcomes that would obtain under these idealized nondiscriminatory conditions, then any departures from this pattern, even departures that have been longstanding, are marks of a discriminatory state of affairs. To the extent that the PC standard sets forth a baseline from which discriminatory harm may be assessed that does not normalize past and present patterns of discrimination, it too offers a worthy perspective on what nondiscrimination requires.

The problem that arises in TJ and in conditions of injustice more broadly is that these two conceptions of discriminatory harm pull in opposite directions. For *even when* key social institutions are marked by injustice, some groups that are socially disadvantaged nonetheless rely on them and moreover may have *come to* rely on them precisely because of past discrimination. In these cases, the BAC is sensitive to the normative interest in safeguarding this reliance. The problem with adopting this standard of discriminatory harm wholesale is that there is also something rather perverse in requiring—*for the sake of nondiscrimination* no less—that an unjust status quo be preserved in order to ensure stability and reliance in this state of affairs. By contrast, the PC standard sets its sights exclusively on some normatively idealized state of affairs. So long as the status quo is unjust, the PC dictates that it must simply be overcome en route to this ideal.

On this way of seeing things, the BAC and PC conceptions of discriminatory harm mark a substantive normative conflict. And contrary to popular portrayals of this conflict, it is not best understood as one between different social groups—say Black and Latino students "versus" Asian students—that are differentially advantaged and disadvantaged by some state of affairs. Rather, we want to suggest that what is at stake here, and in many circumstances marked by injustice, are two different normative goods: that of affirming the forms of agency that have cropped up even in a world marked by unjust obstacles and that of making the world anew in light of past and present injustice. It is important to see that these goods need not track different social groups. The same disadvantaged group might have members who are variously empowered within and constrained by an unjust social system; in fact, the same individual might find themselves in one moment able to successfully maneuver conditions of injustice and in another moment, saddled by those very conditions. It is a fraught fact that the social structure that forms the terrain on which people create their lives is often both an unjust impediment and a condition on which agency is built. This is the profound conflict that gives rise to the opposing baselines against which discriminatory harm is assessed proposed by the BAC and PC: the choice between supporting the very real life plans forged in our actual conditions of past and present injustice and remaking the unjust social structure to be as it should be.

To be sure, there may be circumstances where the two normative concerns do not point in opposing directions and may instead be reconciled with one another. It is possible that in some cases, the best way to support the life plans of disadvantaged social groups now just is to remake unjust social structures; and vice versa, the best way to remake unjust social structures is to start from the real life plans of social

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

groups today and ask what changes to the social structure would facilitate their development and realization. One might even say that this is how we should construct and pursue a robust program of nondiscrimination: by taking our cues from the real-life plans and forms of agency of marginalized social groups today. In pointing out the tension between building anew and supporting the lives built upon existing structures, we do not mean to take issue with this general line of thought. Rather, our point is simply that in an unjust world, many of the lives and plans that people live and make are grounded in those very structures that are part and parcel of the broader unjust social conditions. And this creates a conflict when considering how we should treat those structures, for it suggests that they should neither be straightforwardly dismantled nor continually supported because people have developed lives and plans that rely on them. Selective schools and the prevailing "meritocratic" standards for entrance are an especially illustrative example of this tension. Such schools are, at the same time, bottlenecks *and* ladders to opportunity. Barring broader interventions into the social structure, to change how they function in the first respect is to change how they function in the second. Changes in this context will therefore always be double-edged: a more perfect admissions process (granting such a thing) means that those who had organized their life around the previous process will be left in the lurch.

The practical question at hand is thus, which of the two standards best appreciates the sum total of the relevant normative considerations in our particular circumstances. That this conflict is an inescapable feature of life in an unjust world leads to a perhaps fruitful comparison between the divergence in how the BAC and PC standards treat the unjust status quo and two opposing orientations to theorizing about ethical matters in conditions of injustice. The BAC standard's apparent sensitivity to how social life develops in conditions of injustice—in particular, how conditions of injustice differentially shape the opportunities and choices available to different groups and, in turn, how those groups have come to adapt to and rely on social institutions in different ways—seems to reflect an attentiveness to facts on the ground as they actually stand in our society. By contrast, the PC standard expresses a view of what constitutes a discriminatory effect that is taken from a vision of how goods in an idealized nondiscriminatory society would be distributed. It fixes its gaze on the pattern of outcomes that would obtain under some ideal set of circumstances and does not modify its prescription for the circumstances of our real world.

On this way of seeing things, the opposing BAC and PC standards embody opposing judgments as to the extent to which normative theorizing should incorporate facts about our actual nonideal circumstances in its judgments of what constitutes a discriminatory harm.[33] That discrimination is a normative phenomenon borne out of an unjust world would seem to give an upper hand to the standard that does not idealize away those very features that give rise to discrimination as a social problem. So, that would seem to suggest a point against the PC. However, the conclusion that this orientation thus favors the BAC would be too quick. The BAC's

[33]The divergence in these two perspectives might be described as the difference between the perspective of "non-ideal" and "ideal" theory in political philosophy. This terminology is used to cover a great many distinctions, but here it would track the difference between a standard of what ought to be the case formulated with nonideal circumstances expressly in mind versus a standard of what ought to be the case formulated for an ideal set of circumstances. On this crude division, the BAC would seem to be more in line with the former "non-ideal" orientation and the PC, more in line with the latter "ideal" orientation.

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

status quo baseline does not capture *all* of the morally relevant on-the-ground facts of our actual context. A standard that expresses such a strong presumption in favor of stability and takes policy change alone to indicate a potentially discriminatory effect does not reflect this rich set of considerations and treats as negligible the substantial moral costs of maintaining the status quo. Our nonideal circumstances also give us strong reasons to *destabilize* rather than stabilize our institutions. In the case of TJ, these reasons flow from, among other considerations, the entrenched disadvantage of the Black and Latino communities in the region, their longstanding exclusion from accessing key public goods in the area, and the effects that changing the distribution of the best educational resources could have on mitigating this disadvantage. The normative significance of these facts is lost in the BAC, which, in privileging the status quo ante as it in fact was, rules out all considerations about what ought to be.

VII. Conclusion

U.S. constitutional law appears committed to two propositions. First, that so-called "disparate impact" "standing alone" does not give rise to a violation of equal protection.[34] Second, that invidious intent can render facially neutral policies constitutionally problematic.[35] What we see in *disparate treatment by proxy* cases is that no constitutional violation exists in these cases without *both* invidious intentions and discriminatory harm. This claim bears emphasis. The canonical understanding of equal protection would elevate invidious intention and demote the constitutional significance of discriminatory effect. *Disparate treatment by proxy* cases reveal that this characterization is misleading. Invidious intent (however defined) is not as potent as one might have thought, as standing alone it does not render a facially neutral policy invalid. And, more significantly and perhaps surprisingly, discriminatory effect is necessary as well. In other words, *discriminatory harm* is not constitutionally inert and is as constitutionally significant as invidious intent.

As they both matter constitutionally, it behooves us to have a better understanding of what *discriminatory harm* is and why it is constitutionally important. What account of *discriminatory harm* best explains and justifies its constitutional significance? Furthermore, in plumbing the normative rationales that may underlie the BAC and PC baselines, the idea of discriminatory harm itself emerges as a more intuitive and morally significant notion than it might have appeared at first glance. To be sure, the standards themselves are crude rough-and-ready approximations of the important moral values at stake in a society whose institutional structure both is thoroughly shaped by injustice and nevertheless forms the ground of agency and even uplift for disadvantaged groups. But these shortcomings seem only to bring out as a contrast the genuine moral question of what *should* count as a discriminatory harm in a world like ours. Similarly, the tensions between the BAC and PC explored in this article do not, in our eyes, serve to defeat the plausibility of a notion of discriminatory harm but rather to underscore the complexity of the notion and the challenges inherent in developing an adequate account of it.

Whatever the fate of the BAC and PC baselines in particular, the matter of what really underlies cases of *disparate treatment by proxy* will remain an important issue

[34]Washington v. Davis, 426 U.S. 229, 237 (1976).
[35]*Id.* at 240; *see also* Pers. Adm'r v. Feeney, 442 U.S. 256 (1979).

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press

for theorists of discrimination to grapple with in the years to come. This article has sought to uncover the structure of such claims and their normative underpinnings. The question of what constitutes a discriminatory intent has long occupied a central place in work on discrimination. The time is ripe to now bring another question, that of what constitutes a discriminatory harm, into the fold.

Competing interests. The authors declare none.

Cite this article: Hellman D and Hu L (2025). Disparate Treatment and Discriminatory Harm. *Legal Theory* **31**, 293–315. https://doi.org/10.1017/S1352325226100810

https://doi.org/10.1017/S1352325226100810 Published online by Cambridge University Press